先端データ解析論レポート第二回

学生番号 37-176839 田村浩一郎

2017 年 4 月 23 日

1 宿題 1

ガウスカーネルモデルを実装し，

$$y = \frac{sin(\pi x)}{\pi x} + 0.1x$$

に対する推定を行なった．

そして，正規化パラメタとガウス幅をクロスバリデーションによって決定した．

実装及び出力結果は以下 url の github にある．

`https://github.com/koichiro11/lecture/blob/tamura/ada/ada2/ada_2_tamura.ipynb`

2 宿題 2

線形モデル

$$f_\theta(\mathbf{x}) = \sum_{j=1}^{b} \theta_j \phi_j(x)$$

を用いた L_2 正則化回帰に対する一つ抜き交差確認による二乗誤差は，次式で解析的に計算できることを示す．

$$Loss = \frac{1}{n} ||\tilde{\mathbf{H}}^{\mathbf{-1}}\mathbf{H}\mathbf{y}||^2$$

Proof. $(\mathbf{x_i}, \mathbf{y_i}$ を除いて学習したパラメタは，講義資料より以下のように表せる．

$$\hat{\theta} = (\mathbf{X}_{-\mathbf{i}}^{\mathbf{T}}\mathbf{X}_{-\mathbf{i}} + \lambda\mathbf{I})^{-1}\mathbf{X}_{-\mathbf{i}}^{\mathbf{T}}\mathbf{y}_{-\mathbf{i}}$$

ただし，X_{-i} は $(b-1) \times p$ の行列である．

ここで Sherman-Marrison-Woodbury の公式より，

$$\begin{aligned}
\mathbf{X}_{-\mathbf{i}}^{\mathbf{T}}\mathbf{X}_{-\mathbf{i}} &= (\mathbf{X}^{\mathbf{T}}\mathbf{X} - \mathbf{x}_{\mathbf{i}}^{\mathbf{T}}\mathbf{x}_{\mathbf{i}})^{-1} \\
&= (\mathbf{X}^{\mathbf{T}}\mathbf{X})^{-1} + \frac{(\mathbf{X}^{\mathbf{T}}\mathbf{X})^{-1}\mathbf{x}_{\mathbf{i}}^{\mathbf{T}}\mathbf{x}_{\mathbf{i}}(\mathbf{X}^{\mathbf{T}}\mathbf{X})^{-1}}{1 - (\mathbf{x_i}^{-1}\mathbf{X}^{\mathbf{T}}\mathbf{X})^{-1}\mathbf{x_i}}
\end{aligned}$$

ここで，

$$(\mathbf{X}_{-\mathbf{i}}^{\mathbf{T}}\mathbf{X}_{-\mathbf{i}})^{-1}(\mathbf{X}_{-\mathbf{i}}^{\mathbf{T}}\mathbf{y} - \mathbf{x_i}\mathbf{y_i})$$

を計算することによって，

$$\hat{\theta}_{-\mathbf{i}} = \hat{\theta} - (\mathbf{X_{-i}^T X_{-i}})^{-1}\mathbf{x_i y_i} + \frac{(\mathbf{X^T X})^{-1}\mathbf{x_i^T x_i}(\mathbf{X^T X})^{-1}(\mathbf{X_{-i}^T y} - \mathbf{x_i y_i})}{1 - (\mathbf{x_i}^{-1}\mathbf{X^T X})^{-1}\mathbf{x_i}}$$

上式を整理すると，

$$\hat{\theta}_{-\mathbf{i}} = \hat{\theta} - \frac{(\mathbf{X^T X})^{-1}\mathbf{x_i}(y_i - \mathbf{x_i}^T\hat{\theta})}{1 - h_{ii}}$$

が得られる．

今回対象としている誤差関数 $Loss$ に対して，

$$\begin{aligned} Loss &= \sum_{i=1}^{n}(y_i - \mathbf{x_i}^T\hat{\theta}_{-\mathbf{i}})^2 \\ &= \sum_{i=1}^{n}\frac{y_i - \mathbf{x_i}^T\hat{\theta}}{1 - h_{ii}} \end{aligned}$$

と導ける．そうして，

$$\mathbf{H} = \mathbf{X} - \mathbf{X}(\mathbf{X}^T\mathbf{X} + \lambda\mathbf{X})^{-1}\mathbf{X}^T$$

とした時，h_{ii} は $\tilde{\mathbf{H}}$ の対角成分となっていることに気をつけると，

$$Loss = \frac{1}{n}||\tilde{\mathbf{H}}^{\mathbf{-1}}\mathbf{H}\mathbf{y}||^2$$

□